International Business & Economics Studies
ISSN 2640-9852 (Print) ISSN 2640-9860 (Online)
Vol. 7, No. 4, 2025
www.scholink.org/ojs/index.php/ibes

Original Paper

Internet and Family Entrepreneurship: Is Geographic Distance Still a Problem?

Ke Wang[1*] & Yana Zhou[2]

1. School of Finance, Guangdong University of Finance and Economics;

2. School of Finance and Investment, Guangdong University of Finance

[*] Correspondence: No. 21 Luntou Road, Guangzhou, Guangdong Province, China, 510320.

Email: wk19930115@gdufe.edu.cn. Tel: (+86) 15216402368

Received: May 26, 2025 Accepted: June 28, 2025 Online Published: July 14, 2025

doi:10.22158/ibes.v7n4p33 URL: http://dx.doi.org/10.22158/ibes.v7n4p33

Abstract

This paper explores the value creation mechanism of the Internet from the perspective of family entrepreneurship, using data from the China Household Finance Survey (CHFS). The results show that the use of the Internet alleviates the negative impact of geographic distance on family entrepreneurship, and the effect is more pronounced in rural areas. The Internet reduces the gap between urban and rural areas in family entrepreneurship. Furthermore, the types of information obtained from the Internet by families influence the effects of the Internet. This paper provides evidence that Internet technology can promote entrepreneurial activities at the family level and extends the research on Internet and family entrepreneurship by examining the extent to which the Internet alleviates the negative impact of geographic distance. This supports the notion that the Internet provides a mechanism for economic development of remote and/or underprivileged areas, and therefore represents an important mechanism for promoting rural development.

Keywords

Internet, Family Entrepreneurship, Geographic Distance, Rural Entrepreneurship, Rural Economic Development

JEL Classification: O33, D13, J23

1. Introduction

The value creation effect of entrepreneurial activities has been proven by research. It is well-established that entrepreneurship plays an important role in promoting innovation (Acs & Audretsch, 1990), employment (Acs & Armington, 2004; Fritsch & Weyh, 2006), and economic development

(Wennekers et al., 2005). Family entrepreneurship, which is a part of entrepreneurial activities, may also be a significant factor in economic development, especially in transition economies and less-developed countries. Chinese Premier Li Keqiang put forward the idea of "mass entrepreneurship" in 2014 and regarded it as one of the new engines of China's economy. Since then, what factors affect family entrepreneurship has become one of the focuses of Chinese academic research. Family entrepreneurial activities can directly help poor families to escape from poverty and relieve employment pressure, which is very important for China at this stage, especially considering the goal of achieving a well-off society and getting rid of poverty in an all-round way by 2020. Existing research has shown that geographic distance hinders entrepreneurial activities (Henderson, 2002; Lamb & Sherman, 2010). However, few studies focus on how to deal with the negative effects of geographic distance.

Agglomeration is an important feature of spatial distribution of economic activities (Fujita & Rivera-Batiz, 1988; Fujita et al., 1999) because “centripetal” may mean larger markets (Glaeser et al., 2000; Glaeser, 2007), more knowledge spillovers (Arundel & Geuna, 2004; Acs & Varga, 2005), more social connections (Psaltopoulos et al., 2005; Stam, 2009) and larger labour supply and more diverse skills of workers (Fujita et al., 1999), all of which are important for entrepreneurs. However, starting a business in more densely populated areas means higher land rental costs (Henderson, 1974; Krugman, 1991; Tabuchi, 1998), and potentially higher labour costs (Démurger et al., 2009). For entrepreneurs in remote/rural areas this creates a quandary. On the one hand, a remote location has none of the advantages of urban areas that are important to entrepreneurs at the same time as the average rural entrepreneur cannot afford to move to an urban area or, for social reasons, may not prefer to do so.

In theory at least, the growth and spread of the internet solves this quandary by allowing the benefits of an urban location to be captured in a rural area. In other words, the internet provides a means to overcome the disadvantages of geographic distance. There are a number of existing studies which have investigated the economic effects of the Internet (Vemuri & Siddiqi, 2009; Forero, 2013; Paunov & Rollo, 2016). These studies show that the Internet makes information transmission free from, i.e., independent of, geographic distance and creates a network connected by computers. Meanwhile, the expanded information sharing network creates a virtual platform for commercial activities, which is not bound by space. In line with this, some studies have demonstrated that the information transmission role of the Internet can break down trade barriers created by geographic distance and promote the development of service sector industry (Gassmann & Keupp, 2007; Gnangnon & Iyer, 2018).

For these results to hold requires that the necessary computer-communications infrastructure be available in rural areas. In the case of China, since the start of the 21st century, China's urban and rural Internet penetration rates have increased year by year, and reached 74.6% and 38.4% by the end of 2018, respectively (Figure 1). Although the urban penetration rate is still approximately twice the rural penetration rate, what is significant for our purposes is that the rural rate has grown to almost 40% and for those rural areas which now have internet service, this is significant. In these rural areas, the

Internet has become an important tool for family entrepreneurship, a type of entrepreneurship which tends to be more common in rural areas. According to "China's Taobao Village Research Report (2018)" released by the Ali Research Institute in October 2018, more than 3200 "Taobao villages" (Note 1) have been founded; these are widely distributed in more than 330 districts of 24 provinces in China, and the annual sales volume of these online stores exceeds 220 billion yuan. The villagers of Shaji town in Jiangsu Province, for example, have been widely praised for their wealth creation by relying on network retailing. Shaji town has changed from a backward "garbage town" to an "e-commerce town" with an annual sales volume of more than 7 billion yuan. (Note 2) Although this is just one example, what it demonstrates is two important propositions: 1. the internet provides a means for a small rural area to experience significant economic development and 2. the development in question has been primarily driven by family entrepreneurship.

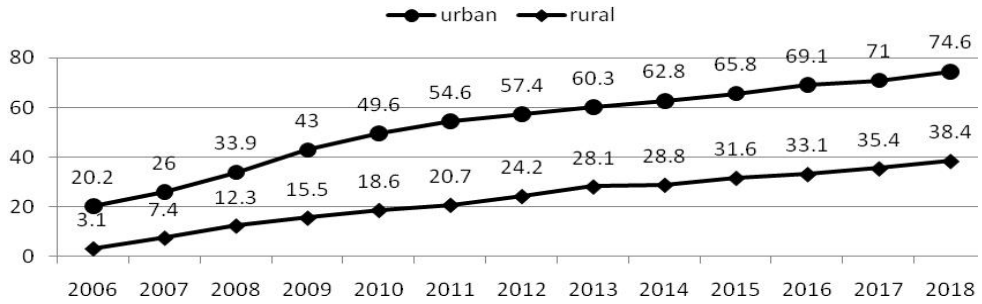


Figure 1. China's Urban and Rural Internet Penetration Rates (2006-2018)

Data source: previous China Statistical Report on Internet Development.

The studies closely related to this paper mainly investigated the information transmission role of the Internet in promoting entrepreneurship. Guillen and Suarez (2001) shows that the Internet is closely related to a country's entrepreneurial environment, using a sample of 141 countries. Wynne et al. (2001) finds that the information acquisition and information transmission functions of the Internet can help potential entrepreneurs to find business opportunities in South Africa's tourism industry. Shen et al. (2018) points out that digital entrepreneurship based on Internet technology will become one of the hottest topics of future research. In addition, existing studies have shown that modern information technology can make economic activities more decentralized (Zaheer and Manrakhan, 2001; Barry, 2011).

On the basis of the above research, this paper focuses on the value creation mechanism of the Internet from the perspective of easing the barriers brought by geographic distance. We try to explore the following questions, using data from the China Household Finance Survey (CHFS). First, is family entrepreneurship constrained by a remote location? Second, if so, can the Internet moderate the relationship between geographic distance and family entrepreneurship? Is the moderating function more pronounced in rural areas? In addition, the paper further investigates whether the Internet reduces

urban-rural differences in family entrepreneurship, and the heterogeneity of different categories of information obtained from the Internet.

This paper innovates in comparison to existing research in three respects. First, it extends the studies about the factors influencing family entrepreneurship by verifying the negative relationship between geographic distance and family entrepreneurship. Second, it enriches the studies about the effects of the Internet on economic activities by discussing family entrepreneurship. Third, this is a pioneering research in the Chinese context, providing empirical evidence of a value creation function of the Internet in transition economies and also guidance on appropriate policies to reduce poverty and promote greater social equality.

The remainder of the paper is organized as follows. Section 2 reviews relevant research and formulates the main hypothesis of the paper. Section 3 describes the data and the empirical methodology. Our initial empirical analysis and further analysis are presented in Section 4 and Section 5. A robust check on the findings is presented in Section 6 and Section 7 offers conclusions.

2. Theory and Hypothesis Development

Densely populated areas are conducive to entrepreneurship for four reasons. First, the densely populated areas usually have high and decentralized market demand, which is conducive to the formation of the scale economic effects of production; second, the positive externality due to knowledge spillover in these areas is beneficial for entrepreneurs; third, people in these areas have more opportunities to build dense social networks, which are crucial to entrepreneurial success; and fourth, workers with particular skills are more likely to be found in a larger population center which has, by definition, more workers to choose from.

From the perspective of market scale, Piore and Sabel (1986) argue that urbanization divides the original market of large-scale standardized production of goods and services into more diversified and smaller markets, which can increase the diversity of urban demand, and is conducive to the establishment of emerging enterprises. Glaeser et al. (2000) and Glaeser (2007) also point out that high population density can promote market demand and consumption. In this way, densely populated areas provide a foundation for entrepreneurial success.

Knowledge accumulation is closely related to entrepreneurial success, while short distance is important for knowledge accumulation (Krugman, 1991). The empirical study of Arundel and Geuna (2004) shows that geographic distance has an important impact on knowledge acquisition and accumulation. However, knowledge spillovers have spatial boundaries (Acs & Audretsch, 1990). Therefore, short geographic distance facilitates entrepreneurship, including family entrepreneurship.

In addition, social ties are essential in entrepreneurial success, and the establishment of social networks needs frequent and intense communication (Granovetter, 1977). Families in remote areas will have greater difficulty in establishing social ties that can help their entrepreneurial activities, compared with families in urban areas (Psaltopoulos et al., 2005), which is one of the reasons why geographic distance

inhibits family entrepreneurship (Stam, 2009).

Finally, the labour supply and the available skills of workers are likely to be larger and more diverse in urban areas and availability of the "right" workers is an important consideration in choosing where to locate a firm. Moreover, this is a two-way street in the sense that once a location is identified as a place where certain skills are in demand, workers who have those skills will move to that location, because the probability of employment is better.

Given these arguments, the first hypothesis to be tested is:

H1: The geographic distance between a family and the city center is negatively correlated with family entrepreneurship.

The superior conditions mentioned above provided by the densely populated zones for entrepreneurs can be looked at in the context of information transmission. Assume initially that there is no internet. Then a number of propositions about information transmission can be made. First, the large-scale market provided by densely populated areas means that information about commodities can be transmitted to consumers more effectively and/or at lower cost.

Second, knowledge is carried and spread through symbols such as words, sounds, images and so on, and expresses specific meanings; thus, knowledge is a kind of special information from the perspective of communication science. Information flow is the key to knowledge transfer (Malecki, 1994). Entrepreneurs close to the center area have stronger/better access to information, so they can enjoy more of the benefits of knowledge spillover than people in remote areas.

Third, social networks, which are a form of information transmission, are an important source of information acquisition and may provide entrepreneurs with more extensive information (Birley, 1985). Therefore, the benefits brought to entrepreneurs in central, more densely-populated areas through social connections, to a large extent reflect the greater ease and convenience of information transmission in such areas. Potential entrepreneurs in remote areas have less access to social networks than those in central areas, which is especially reflected in less social connections outside the family scope (Granovetter, 1977; Song, 2015).

Fourth, the powerful information transmission function of the Internet also enables entrepreneurs to obtain more information about job seekers and select employees from a larger range of labor force. Existing studies have demonstrated the role of the Internet in reducing information asymmetry in the labor market, pointing out that the Internet can shorten the waiting time of unemployed groups (Kuhn & Skurturedu, 2004), improve the labor participation rate of married women (Detting, 2017), and promote the employment growth rate at the regional level (Forman et al., 2012).

The argument here is that information transmission is crucial to entrepreneurial activities. Regions with information transmission advantages tend to be more creative (Anderson, 1985; Johansson, 1987). Lopez (2009) shows that self-entrepreneurship is not only influenced by local information brought by social networks, but also by information transmission worldwide. However, information transmission is limited, or made more costly, by spatial distance, which contributes to a regional inequality in

individual communication (Malecki, 1994). But now, let's allow the internet to be present. Location, to a large degree, now becomes irrelevant to information transmission. Information can now be accessed from anywhere, stored anywhere, accessed at any time, accessed in the same time, saved and retransmitted, at approximately the same cost and in the same quantity. In other words, the inequality between urban and rural areas in access to information and ease and cost of information transmission is eliminated by the advent of the internet. As a revolutionary tool of information transmission, the Internet has changed the characteristics of information transmission, and has broken the limitations of time, space, and content of information transmission. The advantages of information transmission through the Internet can be reflected in three aspects: the Internet can complete the instantaneous transmission of information over a long distance; the Internet makes everyone become a potential sender of information, rather than just the receiver of information; and the information transmitted by the Internet can be presented in more ways, not just words and pictures.

Basically, what is being argued here is that the negative effects of geographic distance on entrepreneurship are effectively eliminated or at least largely redressed once we allow the internet to exist. With the Internet, entrepreneurs can release information on their products to a large number of potential consumers through Internet E-commerce and other specialized Internet platforms, thus solving the problem of limited market scale in remote areas; second, in terms of knowledge acquisition, the Internet can not only provide various kinds of free knowledge and information, but also can provide paid knowledge services, thus bridging the knowledge gap of potential entrepreneurs in remote areas; third, social networking tools, such as Wechat and Weibo, make it easier for entrepreneurs in remote areas to establish social network or to tie into them; forth, The Internet can provide entrepreneurs with a wider range of choices in terms of hiring employees.

Although the above analysis shows the support role of the Internet for family entrepreneurship from the perspective of alleviating the distance factor, it is obvious that the Internet is not omnipotent. It is worth noting that the Internet cannot solve the transportation cost problem caused by distance. However, the following two realities from China may ease that concern. First, China has a developed express network. A large number of online shopping activities carried out by residents all over the country have given birth to China's very developed express delivery industry. Goods can usually be transported at a lower cost and in a shorter time. More importantly, the permeability of China's express network is very strong, and many rural families can enjoy almost the same express service as urban areas within a short distance from home. Second, China has been actively promoting road construction in rural areas. The high permeability of express network depends on the high permeability of road. One fact is that almost all villages in China have been connected to roads, which has greatly alleviated the transportation problems in remote areas.

To sum up, we believe that the changes brought about by the Internet for information transmission has enabled potential entrepreneurs to have more powerful information acquisition and transmission capabilities, thus reducing the problems of small market scale, inconvenient knowledge acquisition,

fewer social connections and lack of labour caused by geographic distance, and promoting family entrepreneurship in remote areas. This generates the second hypothesis for testing:

H2: The Internet can alleviate the negative impact of geographic distance on family entrepreneurship.

Rural areas are generally considered to be far away from the central area of human activities, and faced with more severe conditions in production and distribution and a larger information gap, compared with urban areas. For example, rural areas rely more on internal connections when they obtain goods, labor services, information, etc., and this weakens their social connections and business connections with the outside world (Dabson, 2001). The remote location and narrow space make it difficult for rural areas to develop economies of scale, which leads to high prices and low demand for goods and services. In addition, the imperfect transportation infrastructure makes it more difficult, if not impossible, to connect with the external market, resulting in less knowledge spillovers (Henderson, 2002; Lamb & Sherman, 2010). The problem of poor information transmission caused by remote geographic location is more likely to be concentrated in rural areas, and this problem is more difficult to be alleviated by other means due to the backward technology and transportation infrastructure. Therefore, we expect that the moderating effect of the Internet will be more pronounced in rural areas, thus leading to our third hypothesis:

H3: the moderating effect of the Internet on the relationship between geographic distance and family entrepreneurship is more pronounced in rural areas.

3. Research Design

3.1 Sample and Data Sources

This paper uses data from the China Household Finance Survey (CHFS), which is a national household survey conducted by the China Household Finance Survey and Research Center of Southwestern University of Finance and Economics. Data from CHFS for 2011 is used. (Note 3) The final sample includes information on 8438 families in 25 provinces. Our research focuses on the relevant information at the family level, such as work, house, Internet usage, and so on. We also control some personal factors. We only keep sample information where the answer to the question "what is the relationship with the interviewee?" is "myself"; this personal information is then matched with the relevant information at the family level.

3.2 Model Specification and Variable Definitions

To test our hypotheses, we use the following regression model:

$$selfemp = \beta_0 + \boldsymbol{\beta_1} \times DIS + \boldsymbol{\beta}_2 \times CONTROL + \varepsilon \quad (1)$$

$$selfemp = \beta_0 + \boldsymbol{\beta_1} \times DIS + \beta_2 \times internet + \boldsymbol{\beta}_3 \times internet \times DIS + \boldsymbol{\beta}_4 \times CONTROL + \varepsilon \quad (2)$$

where the dependent variable (*selfemp*) is a dummy variable equal to 1 if the answer to “Is your family engaged in industrial and commercial production?” is “yes”, and equal to 0 otherwise. Our key independent variable is geographic distance (*DIS*). To estimate the distance between the family and the

city center, based on respondents' estimates of the time spent travelling from their family residence or their own houses to the city center and the types of transportation used, we construct the following three indicators as proxy variables. The actual distance between the family residence and the city center (*dis_live*), the average distance between family-owned houses and the city center (*dis_aver*), and the minimum distance between family-owned houses and the city center (*dis_min*). These all reflect the remoteness of the family's location. The larger the value, the farther away the family's location is.

Internet is an indicator reflecting the Internet usage of families, which is a dummy variable equals to 1 if the answer to "what is the main source of your information?" is "Internet", and 0 otherwise. It should be noted that this indicator not only refers to the use of computers to access the Internet, but also covers the use of any device (mobile phones, tablet computers, etc.) that provides access to the Internet.

Following other studies, our control variables include the number of family members (*members*), the number of children (*kidnum*) under 16 years old, the number of the head of household and his/her spouse's brothers and sisters (*bsnum*), whether there are elderly people in the household (*elder*, 65 years old and above), and the total assets of the family (*asset_hh*, the logarithm of the total assets). We also control personal factors using gender (*gender*), risk preference (*risk*), age (*age*), the square of age (*agesqu*), marriage (*marriage*), education (*edu*), minority nationality or not (*minor*), members of the Communist Party of China or not (*comm*), local resident or not (*acc_local*), belonging to rural registered population or not (*acc_agri*). The control variables at the regional level include the average per capita GDP for the period 2007-2011 (*gdp_pe*, the logarithm of GDP/capita), road density (*roadde*), and regional business atmosphere (*atmosp*). In addition, we add family-level urban-rural factor (*rural*) to the control variables, in order to examine the urban-rural differences in the role of the Internet.

3.3 Summary Statistics

We winsorize, i.e., reduce the influence of spurious outliers of the three distance variables (*dis_live, dis_aver, dis_min*) and family total assets (*asset_hh*) at the levels of 1% and 99%. Table 1 shows that about 13.3% of the families are engaged in industrial and commercial production, and about 25.7% use the Internet to obtain information. Additionally, the average of the three distance indicators is about 20km. (Note 4)

Table 1. Summary Statistics

Variables	Definitions	Num	Mean	Std_dev
selfemp	A dummy variable equal to 1 if the answer to "Is your family engaged in industrial and commercial production?" is "yes", and 0 otherwise.	8436	0.133	0.340
internet	A dummy variable equals to 1 if the answer to "what is the main source of your information?" is "Internet", and 0 otherwise.	8422	0.257	0.437

Variables	Definitions	Num	Mean	Std_dev
dis_live	The distance between the family residence and the city center.	8362	19.47	20.63
dis_aver	The average distance between family-owned houses and the city center.	7567	20.98	22.11
dis_min	The minimum distance between family-owned houses and the city center.	7590	20.38	22.15
rural	A dummy variable equal to 1 if the family is located in rural areas, and 0 otherwise.	8438	0.384	0.486
member	The number of family members.	8438	3.475	1.548
kidnum	The number of children (under 16 years old).	8438	0.549	0.753
bsnum	The number of the head of household and his/her spouse's brothers and sisters.	8281	5.782	3.399
elder	A dummy variable equal to 1 if there are elderly people (65 years old and above) in the family, and 0 otherwise.	8436	0.272	0.445
asset_hh	The total assets of the family (thousand RMB).	8438	488.26	1.937
gender	A dummy variable equal to 1 if gender is male, and 0 otherwise.	8383	0.538	0.499
risk	Risk preference of the respondents, equal to -1 if risk preference is low, 0 if risk preference is average, 1 if risk preference is high. (Note 5)	8295	-0.472	0.721
age	Age of the respondent.	8383	48.86	14.45
marriage	A dummy variable equal to 1 if the respondent is married and 0 otherwise.	8297	0.864	0.343
edu	Years of education of the respondent.	8306	9.169	4.364
minor	A dummy variable equal to 1 if the respondent is of a ethnic minorities, and 0 otherwise.	8308	0.032	0.177
comm	A dummy variable equal to 1 if the respondent is a member of the Communist Party of China, and 0 otherwise.	8305	0.149	0.356
acc_local	A dummy variable equal to 1 if the respondent is a local resident and 0 otherwise.	8301	0.922	0.268
acc_agri	A dummy variable equal to 1 if the respondent is an agricultural resident and 0 otherwise.	8364	0.528	0.499
gdp_pe	The average per capita GDP from 2007 to 2011 (the logarithm of GDP/capita).	8438	10.32	0.474
roadde	Provincial highway mileage (km) / provincial area (km^2) in	8438	1.102	0.485

Variables	Definitions	Num	Mean	Std_dev
	2011.			
atmosp	The number of employees in private enterprises and individual enterprises/the provincial population*100 in 2011.	8438	15.310	6.955

In terms of control variables, about 38.4% of the sample families come from rural areas, with an average of 3.475 members and 0.549 children. The total number of brothers and sisters of the respondents and their spouses is 5.782; 27.2% of the sample families have elderly members. The average total assets of the sample families is close to 490000 RMB, 53.8% of the respondents are men, 86.4% are married, the average age is 49 years old, and the average number of years of education is more than 9 years. The number of people who belong to the risk aversion type is more than those belonging to the risk-preference type. In addition, only 3.2% of the total samples are minority families, 14.9% are CPC members. 92.2% and 52.8% of the total samples are local residents and agricultural residents, respectively. (Note 6)

3.4 Grouping Test of Variables

Next, we take the dependent variable (*selfemp*) as the grouping variable and carry out the single variable statistical analysis and t-test to the main variable, so as to initially show the relationship between the main variables and the dependent variable. As can be seen from the results in Table 2, five main variables showed significant differences at the 1% level in the two groups. First, 38.6% of the families in the entrepreneurial sample use the Internet to obtain information, 14.9% higher than that in the non-entrepreneurial sample. Second, for the three indicators of residential distance, *dis_live*, *dis_aver*, and *dis_min*, are 14.97km, 17.17km and 16.15km, respectively for the entgrepreneurial sample, less than that in the non-entrepreneurial sample (20.16km, 21.57km and 21.04km, respectively). In addition, only 29% of the entrepreneurial families come from rural areas, which is significantly lower than 39.9% of the non-entrepreneurial group at the level of 1%, which reflects the urban-rural differences of family entrepreneurship. In general, the above results show that there is a strong negative relationship between geographic distance and family entrepreneurship.

Table 2. Grouping Test of Variables

Variables	Sample	*selfemp*=1		*selfemp*=0		T test	
		Sub-sample	Mean	Sub-sample	Mean	Diff	T value
internet	8422	1122	0.386	7299	0.237	0.149	10.71***
dis_live	8362	1116	14.97	7245	20.16	-5.185	-7.842***
dis_aver	7567	1013	17.17	6554	21.57	-4.400	-5.906***
dis_min	7590	1018	16.15	6572	21.04	-4.890	-6.574***
rural	8438	1124	0.290	7312	0.399	-0.109	-7.015***

Note: * * *, * *, * are significant at the level of 1%, 5%, and 10%, respectively. The sum of the two sets of sub-samples may be different from the total number of samples, since the dependent variable

(*selfemp*) has two missing values.

4. Empirical Results

4.1 Remoteness and Family Entrepreneurship

In order to test hypothesis 1, we carry out a probit regression using model (1), and the results are shown in Table 3. The first three columns are empirical results without control variables, while the last three columns are results with all the control variables. It can be seen that the coefficients of all three distance indicators are negative and significant at the 1% level, indicating that the more remote the family is, the less likely it is to initiate entrepreneurial activities. In addition, the coefficient of rural is also negative and significant at the 1% level, indicating that the possibility of family entrepreneurship in rural areas is less than urban areas. Therefore, we initially believe that geographic distance hinders family entrepreneurship.

Table 3. Remoteness and Family Entrepreneurship

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	*selfemp*	*selfemp*	*selfemp*	*selfemp*	*selfemp*	*selfemp*
dis_live	-0.0078***			-0.0072***		
	(0.0010)			(0.0012)		
dis_aver		-0.0056***			-0.0054***	
		(0.0009)			(0.0011)	
dis_min			-0.0062***			-0.0056***
			(0.0010)			(0.0011)
rural				-0.2948***	-0.3028***	-0.2966***
				(0.0544)	(0.0558)	(0.0558)
member				0.1020***	0.0965***	0.0962***
				(0.0168)	(0.0176)	(0.0175)
kidnum				-0.0095	-0.0071	-0.0064
				(0.0305)	(0.0318)	(0.0318)
bsnum				0.0021	-0.0020	-0.0020
				(0.0068)	(0.0070)	(0.0070)
elder				-0.1050**	-0.0860	-0.0830
				(0.0521)	(0.0540)	(0.0539)
asset_hh				0.1207***	0.1613***	0.1597***
				(0.0145)	(0.0185)	(0.0184)
gender				0.1193***	0.1360***	0.1360***
				(0.0392)	(0.0411)	(0.0411)

risk				0.1320***	0.1175***	0.1189***
				(0.0258)	(0.0270)	(0.0270)
edu				-0.0007	-0.0048	-0.0044
				(0.0065)	(0.0069)	(0.0069)
minor				-0.1447	-0.2073	-0.2066
				(0.1221)	(0.1340)	(0.1341)
age				0.0106	0.0078	0.0082
				(0.0101)	(0.0105)	(0.0104)
agesqu				-0.0023**	-0.0021*	-0.0021**
				(0.0010)	(0.0011)	(0.0011)
marrige				-0.0198	-0.0002	-0.0015
				(0.0655)	(0.0702)	(0.0699)
comm				-0.2717***	-0.2378***	-0.2468***
				(0.0636)	(0.0661)	(0.0661)
acc_agri				0.3211***	0.3563***	0.3544***
				(0.0536)	(0.0573)	(0.0572)
acc_local				-0.2236***	-0.2098***	-0.2231***
				(0.0685)	(0.0768)	(0.0762)
gdp_pe				-0.4911***	-0.5178***	-0.5210***
				(0.0855)	(0.0914)	(0.0915)
roadde				-0.0073	-0.0121	-0.0066
				(0.0479)	(0.0499)	(0.0499)
atmosp				0.0278***	0.0287***	0.0286***
				(0.0051)	(0.0054)	(0.0054)
_cons	-0.9723***	-0.9992***	-0.9900***	2.0723**	1.8854**	1.9330**
	(0.0236)	(0.0251)	(0.0249)	(0.8313)	(0.8847)	(0.8847)
N	8361	7567	7590	8063	7309	7330
pseudo R^2	0.011	0.007	0.008	0.103	0.104	0.105

Note: * * *, * *, * are significant at the level of 1%, 5%, and 10%, respectively. The same below.

As for the control variables at the level of family, the total number of families and the total assets of families are positively related with family entrepreneurship. In terms of personal characteristics, males, families with more risk preference interviewees, and agricultural residents are more likely to carry out family entrepreneurship; age and the probability of entrepreneurship show an inverted U-shaped relationship; families with communists and local residents are less likely to initiate entrepreneurial activities. In addition, family entrepreneurship is more likely to occur in areas with lower GDP per capita and a better business atmosphere.

4.2 Endogeneity Problems

Reverse causality may affect the robustness of previous conclusions, that is, entrepreneurs are more likely to choose a location closer to the city center. To solve this problem may still draw useful conclusions, although this kind of reverse causality just proves that the remote geographical location hinders the family entrepreneurship: it means that there is a significant and robust relationship between geographic distance and family entrepreneurship, if the conclusion is still tenable after the elimination of the concern of reverse causality. In this part, we will use a series of methods to investigate this endogeneity problem.

4.2.1 Sample Adjustment

We investigate whether the negative correlation between distance and family entrepreneurship, and the difference between urban and rural areas in Table 3 are caused by population shift. First, we explore whether the migration of rural population to cities (named rural migration) has an impact on the previous conclusions, by removing the sample that belongs to agricultural residents who do not live in rural areas from the total sample. Second, we explore the potential impact of household registration migration (named registration migration) by removing the samples that are not registered in local areas. Third, we further investigate the impact of population shift by removing the samples whose residences are not their own houses (named house renting).

The results are shown in Table 4. Column (1) is the results after excluding the sample of rural migration, and dropping the control variable *acc_agri* in case of serious multicollinearity problems. Columns (2) and (3) are the results after excluding the samples of registration migration and house renting. The results show that the negative correlation between distance and family entrepreneurship is still significant at the 1% level. The coefficients of *rural* are consistent with those in Table 3, except column (1), which shows that the phenomenon of rural migration explains the urban-rural difference of family entrepreneurship to a certain extent.

Table 4. Sample adjustment.

	(1)	(2)	(3)
Variables	*selfemp*	*selfemp*	*selfemp*
Samples	Rural migration removed	Registration migration removed	House renting removed
dis_live	-0.0071***	-0.0065***	-0.0055***
	(0.0013)	(0.0012)	(0.0013)
rural	0.0033	-0.3525***	-0.2752***
	(0.0616)	(0.0581)	(0.0603)
control variables	yes	yes	yes
_cons	3.1163***	2.8944***	1.7459*

	(0.9719)	(0.8814)	(0.9234)
N	6559	7437	6793
pseudo R^2	0.091	0.100	0.099

Note. We only report the results of using the variable of *dis_live* for brevity. The results of using *dis_aver* and *dis_min* are basically the same. For the sake of brevity, we omit the regression results of control variables.

4.2.2 Instrumental Variable

The indicators of distance are calculated by multiplying the time by the estimated speed of the corresponding vehicle, reflecting the horizontal distance. In the previous sections, we have controlled the regional per capita GDP, road network density, business atmosphere and other variables to limit the impact of terrain factors on family entrepreneurship through other channels such as economic development, traffic convenience and so on. If the terrain of a region is relatively steep, the same horizontal displacement means a larger actual distance. Therefore, in order to better solve this problem, we introduce an instrumental variable (*stdeleva*), which reflects the degree of terrain roughness to test the hypothesis.

The results are shown in Table 5, in which column (1) is the regression results of the first stage of two-step ivprobit, column (2) is the regression results of the second stage, and column (3) presents the simplified regression results. In the first stage, the regression results show that the instrumental variable stdeleva and the independent variable *dis_live* have a significant and positive correlation at the 1% level, indicating that the instrumental variable met the correlation requirements. The results of the second stage show that geographic distance has a significant and negative relationship with family entrepreneurship at the 1% level, which is consistent with the results in column (3).

Table 5. Instrumental Variable Regression

	(1)	(2)	(3)
Variables	*dis_live*	*selfemp*	*selfemp*
IV: *stdeleva*	1st stage	2nd stage	Simplification
stdeleva	0.0065***		-0.0002***
	(0.0007)		(0.0001)
dis_live		-0.0327***	
		(0.0101)	
rural	15.4081***	0.0995	-0.4026***
	(0.5492)	(0.1646)	(0.0523)
control variables	yes	yes	yes
_cons	-30.4511***	2.0252**	2.8654***
	(9.1283)	(0.8769)	(0.8498)

N	8063	8063	8129
R^2 or pseudo R^2	0.218	-	0.098

Note. We only report the results of using the variable *dis_live* for brevity. The results of using *dis_aver* and *dis_min* are basically the same. For the sake of brevity, we omit the regression results of control variables.

4.3 The Role of Internet

4.3.1 Full sample

According to the previous analysis, the use of Internet may ease the obstacles of geographic distance to potential entrepreneurs, and alleviate the negative effect of distance on family entrepreneurship. We use model 2 to test the hypothesis, by introducing the interaction of Internet and distance. In addition, we group the full sample into rural families and urban families to investigate the different role of Internet in rural and urban areas.

Table 6 shows the results. Columns (1) and (2) use *dis_live* as the indicator of remoteness of the family, columns (3) and (4) use *dis_aver*, while the last two columns use *dis_min*. The odd columns are the results for urban families (*rural* = 0), and the even columns are the results for rural families (*rural* = 1). The results show that the coefficients of the interaction (*internet*DIS*) are positive and significant at the 5% level, indicating that the Internet can significantly alleviate the negative effect of distance on family entrepreneurship. However, the role of the Internet is not significant in urban families. In general, the results show that the role of the Internet is reflected in rural areas, but not in urban areas.

Table 6. The Role of Internet: Grouping Test of Rural and Urban Areas

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	*selfemp*	*selfemp*	*selfemp*	*selfemp*	*selfemp*	*selfemp*
DIS=	*dis_live*	*dis_live*	*dis_aver*	*dis_aver*	*dis_min*	*dis_min*
Sample	Urban areas	Rural areas	Urban areas	Rural areas	Urban areas	Rural areas
DIS	-0.0032	-0.0059***	-0.0022	-0.0059***	-0.0022	-0.0058***
	(0.0023)	(0.0018)	(0.0019)	(0.0017)	(0.0020)	(0.0017)
*internet*DIS*	-0.0028	0.0121**	-0.0024	0.0128***	-0.0034	0.0137***
	(0.0032)	(0.0054)	(0.0029)	(0.0049)	(0.0030)	(0.0048)
internet	0.1660**	0.2305	0.1362*	0.1746	0.1454**	0.1661
	(0.0687)	(0.1624)	(0.0733)	(0.1613)	(0.0723)	(0.1570)
control variables	yes	yes	yes	yes	yes	yes
_cons	2.6769***	1.3802	2.5774**	1.2699	2.6058**	1.1882
	(1.0101)	(1.7186)	(1.0807)	(1.8040)	(1.0797)	(1.8039)

N	4987	3070	4360	2943	4375	2949
pseudo R^2	0.120	0.121	0.118	0.123	0.119	0.124

Note. For the sake of brevity, we omit the regression results of control variables.

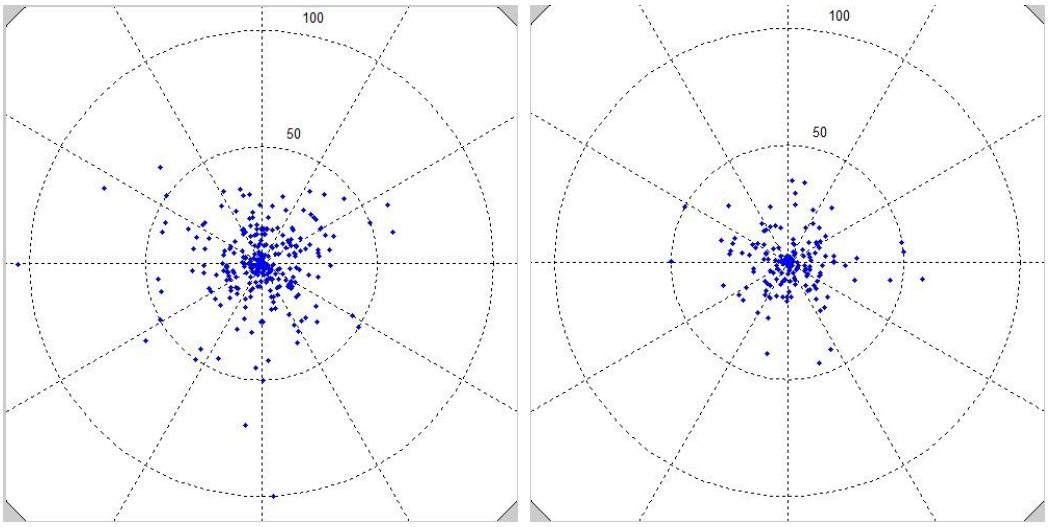
4.3.2 PSM test

The above results can prove the moderating effect of the Internet to some extent. A possible concern is that families close to the city center have more opportunities to use the Internet, so the previous results may be caused only by distance rather than the differences in Internet use. Therefore, we divide the sample into two groups according to whether they use Internet or not, and match them with one-to-one calipers at the 1% level according to the value of distance, so as to ensure that there is no significant difference in the two subsamples after matching. As we can see from Table 7, there is no significant difference between the two groups after matching (take *dis_live* as an example).

Table 7. Grouping Tests of Geographic Distance before and After Matching

Variables	Stage	*internet*=1		*internet*=0		T test	
		Sample	Mean	Sample	Mean	Diff	T value
dis_live	Before matching	2147	14.002	6200	21.347	-7.345	-14.394***
	After matching	2058	14.022	2053	14.053	-0.0313	-0.0566

We use the matched samples and draw a simulated spatial distribution map of family entrepreneurship with the help of polar coordinates, in order to display the moderating role of Internet more clearly and more intuitively (see Figure 2). Each point in the map represents an entrepreneurial family, and the center of the circle represents the central area of the city. The left one is the simulated map with families using Internet, while the right one represents the families who do not use Internet. It can be seen that in the two groups with no significant difference between the numbers in the samples and the average distance from the center, there are more entrepreneurial families in the left map and they are farther away from the center, while there are fewer entrepreneurial families in the right map and they are concentrated within 50km from the center. Table 8 further confirms that the number of entrepreneurial families who use the Internet is significantly larger than those who do not use the Internet, and the former are farther away from the center on average. These results indicate that the Internet can break the barriers brought by distance, and promote family entrepreneurship.



Families who use the Internet Families who do not use the Internet

Figure 2. Simulated Spatial Distribution map of Family Entrepreneurship

Table 8. Grouping Tests of Geographic Distance Within Entrepreneurial Families

Variables	*internet*=1 and *selfemp*=1		*internet*=0 and *selfemp*=1		T test	
	Sample	Mean	Sample	Mean	Diff	T value
dis_live	412	12.449	242	10.347	2.102	1.994**

Next, we use the matched samples to repeat the regression in Table 6 again, and the results are shown in Table 9. Similarly, the first two columns use *dis_live* as a measure of the remoteness of family location, the middle two columns use *dis_aver*, and the last two columns use *dis_min*. The odd columns are the results for urban families (*rural* = 0), while the even columns are the results for rural families (*rural* = 1).

The results in Table 9 further show that the Internet has reduced, if not eliminated, the inhibition of geographic distance on family entrepreneurship. The coefficients of the interaction (*internet* * *DIS*) in the even columns are positive and significant at the 10% level, and the effect is more pronounced in rural areas. In addition, consistent with the results in Table 6, the role of the Internet is not significant in urban areas.

Table 9. PSM Test

Variables	(1) *selfemp*	(2) *selfemp*	(3) *selfemp*	(4) *selfemp*	(5) *selfemp*	(6) *selfemp*
DIS=	*dis_live*	*dis_live*	*dis_aver*	*dis_aver*	*dis_min*	*dis_min*
Sample	Urban	Rural	Urban	Rural	Urban	Rural

	areas	areas	areas	areas	areas	areas
DIS	-0.0140***	-0.0102**	-0.0028	-0.0103***	-0.0034	-0.0034
	(0.0044)	(0.0042)	(0.0038)	(0.0036)	(0.0045)	(0.0040)
*internet*DIS*	0.0082	0.0167**	-0.0033	0.0183***	-0.0027	0.0123*
	(0.0050)	(0.0069)	(0.0044)	(0.0060)	(0.0050)	(0.0064)
internet	0.1196	0.1062	0.1187	0.0255	0.1185	0.4482**
	(0.0857)	(0.2112)	(0.0946)	(0.2044)	(0.0934)	(0.2011)
control variables	yes	yes	yes	yes	yes	yes
_cons	2.2621*	1.8010	2.4471*	1.0103	1.7020	0.4228
	(1.2685)	(2.9493)	(1.3644)	(3.2940)	(1.3560)	(2.7701)
N	3090	902	2585	822	2624	840
pseudo R^2	0.131	0.177	0.125	0.147	0.142	0.204

Note: For the sake of brevity, we omit the regression results of control variables.

4.4 Summary

In general, the above empirical results verify the hypotheses of this paper. First, the distance between family residence and the city center are negatively related to family entrepreneurship, and the effect is most pronounced in rural areas. The results are robust after considering the potential impacts of rural migration, registration migration and families' house renting, and using the regional elevation standard deviation as an instrumental variable for an endogeneity test. Second, the Internet can effectively compensate for the obstacles caused by distance to family entrepreneurship in rural areas, while the effect is not significant in urban families. The results confirm the value creation mechanism of the Internet alleviating the negative effect of geographic distance on family entrepreneurship.

5. Further Analysis

5.1 Does the Internet Reduce the Gap between Urban and Rural Areas in Entrepreneurial Activities?

This section focuses on the urban-rural differences between the Internet and family entrepreneurship and discusses whether the use of the Internet can reduce the gap between urban and rural areas in entrepreneurial activities. For this reason, we introduce the interaction of Internet usage and rural areas (*internet*rural*) and observe the sign and significance of this term. The results are shown in Table 10. It can be seen that *rural* has a significant and negative correlation with the dependent variable at the 1% level, indicating that rural families are less likely to carry out entrepreneurial activities than urban ones. The interaction (*internet*rural*) has a positive and significant correlation with the dependent variable at the 1% level in all three equations, showing the moderating effect of the Internet. The results show that Internet has indeed reduced the urban-rural gap in family entrepreneurship.

Table 10. Does Internet Reduce the Gap between Urban and Rural Areas in Entrepreneurial Activities?

	(1)	(2)	(3)
Variables	*selfemp*	*selfemp*	*selfemp*
DIS=	*dis_live*	*dis_aver*	*dis_min*
rural	-0.4024***	-0.4040***	-0.3964***
	(0.0592)	(0.0606)	(0.0606)
*internet*rural*	0.5431***	0.5282***	0.5183***
	(0.1086)	(0.1118)	(0.1120)
internet	0.0951*	0.0728	0.0726
	(0.0556)	(0.0586)	(0.0585)
DIS	-0.0065***	-0.0050***	-0.0051***
	(0.0012)	(0.0011)	(0.0011)
control variables	yes	yes	yes
_cons	2.3993***	2.2147**	2.2541**
	(0.8398)	(0.8934)	(0.8931)
N	8057	7303	7324
pseudo R^2	0.109	0.110	0.111

Note. For the sake of brevity, we omit the regression results of control variables.

5.2 The Impacts of Heterogeneous Information Categories

In this part, we will discuss whether the categories of information obtained by families through the Internet influence the role of the Internet. In addition to asking the interviewees which channels to use to obtain information, the survey also collected the categories of the information obtained, including the specific categories of economic, social, entertainment, and so on. We construct three dummy variables, including *int_eco*, *int_soc* and *int_ent*, which represent the situation of using Internet to obtain economic information, social information and entertainment information, respectively. The variable equals 1 if the family uses Internet to obtain the corresponding type of information, and 0 otherwise. We use these three variables to replace the variable Internet in columns (1) and (2) of Table 6, and use *dis_live* as the indicator to measure the remoteness of the family. The results are shown in Table 11. The first two columns examine the role of using the Internet to obtain economic information, the middle two columns examine social information, and the last two columns examine entertainment information. The odd and even columns investigate the samples of urban and rural families, respectively.

Table 11. The Impacts of Heterogeneous Information Categories

Variables	(1) *selfemp*	(2) *selfemp*	(3) *selfemp*	(4) *selfemp*	(5) *selfemp*	(6) *selfemp*
Sample	Urban areas	Rural areas	Urban areas	Rural areas	Urban areas	Rural areas
Information categories	Economic	Economic	Social	Social	Entertain	entertain
dis_live	-0.0035*	-0.0057***	-0.0040**	-0.0060***	-0.0029	-0.0057***
	(0.0021)	(0.0017)	(0.0020)	(0.0017)	(0.0019)	(0.0017)
*int_eco*dis_live*	-0.0037	0.0118*				
	(0.0032)	(0.0062)				
int_eco	0.1364**	0.2805				
	(0.0690)	(0.1789)				
*int_soc*dis_live*			-0.0023	0.0134**		
			(0.0036)	(0.0066)		
int_soc			0.0412	-0.1682		
			(0.0717)	(0.2026)		
*int_ent*dis_live*					-0.0090**	0.0119
					(0.0046)	(0.0077)
int_ent					0.1276	0.0264
					(0.0794)	(0.2134)
control variables	yes	yes	yes	yes	yes	yes
_cons	2.7202***	1.4296	2.6027***	1.3839	2.5560**	1.4384
	(1.0126)	(1.7186)	(1.0066)	(1.7110)	(1.0060)	(1.7077)
N	4983	3063	4983	3063	4983	3063
pseudo R^2	0.120	0.119	0.119	0.113	0.120	0.113

Note. We only report the results of using the variable of *dis_live* for brevity. The results of using *dis_aver* and *dis_min* are basically the same. For the sake of brevity, we omit the regression results of control variables.

Columns (1) and (2) show that the interaction *int_eco*dis_live* is significant in rural families at the 10% level, while the effect is not significant in urban families, when using Internet to obtain economic information. The results are consistent with the role of social information in columns (3) and (4), showing that economic information and social information obtained from the Internet can alleviate the negative impact of geographic distance on family entrepreneurship. Columns (5) and (6) indicate that entertainment information can't alleviate the negative impact of geographic distance, and it even exacerbates the negative impact, which may means that, to some extent, paying attention to

entertainment information takes up the time that families could be used to promote products, acquire knowledge, accumulate contacts and search labour force which is more vital for remote families. The results show the different impacts of heterogeneous information categories.

6. Robustness Tests

6.1 Adjustment of Distance Indicator

Estimating the distance between the family and its city center is an important basic work for the empirical analysis in this paper. The distortion of this index will have a great impact on the conclusions of this paper. In order to eliminate this concern, we use the following two methods to reconstruct the distance index.

6.1.1 Adjustment of the Speed of Vehicles

First of all, we drop the samples whose answer to “the types of vehicles to the city center” is “other”, and only consider the several typical types of vehicles given in the questionnaire, namely walking, cycling, motorcycle or electric vehicle, bus, subway, and automobile. In this way, we can effectively avoid the impact of the deviation of respondents when they describe their own vehicles, and estimate the distance more accurately using the standardized types of vehicles. Secondly, considering the waiting time and stop time when using public transportation, we reduce the assignment of the speed of "bus" and "subway" from the original 30km / h and 70km / h to 20km / h and 50km / h respectively, so as to make the assignment of speed more consistent with the actual travel speed. Through the above methods, we construct three new distance indicators, namely, the distance between the family residence and the city center (*addis_live*), the average distance between the self-owned house and the city center (*addis_aver*), and the minimum distance between the self-owned house and the city center (*addis_min*). We then use these indicators to test the hypotheses. The results in panel A of Table 12 are consistent with the previous sections.

6.1.2 Replace Distance with Time Spent

We also replace distance with time spent from family residence to city center to proxy the remoteness of the family. Although this measure ignores the differences among vehicles, it reflects the convenience of families to go to the city center from the perspective of time cost more intuitively. Accordingly, we construct the time spent from the residence of the family to the city center (*ti_live*), the average time spent from each family's house to the city center (*ti_aver*), and the minimum time spent from each family's house to the city center (*ti_min*), respectively. We then use these indicators to test the hypotheses. The results in panel B of Table 12 are consistent with the previous sections.

Table 12. Adjustment of Distance Indicator

	(1)	(2)	(3)	(4)	(5)	(6)
	selfemp	*selfemp*	*selfemp*	*selfemp*	*selfemp*	*selfemp*

Panel A: Adjustment of the speed of vehicles

Sample	Urban areas	Rural areas	Urban areas	Rural areas	Urban areas	Rural areas
DIS=	*addis_live*	*addis_live*	*addis_aver*	*addis_aver*	*addis_min*	*addis_min*
DIS	-0.0040	-0.0064***	-0.0020	-0.0062***	-0.0020	-0.0063***
	(0.0030)	(0.0025)	(0.0027)	(0.0024)	(0.0027)	(0.0024)
*internet*DIS*	-0.0029	0.0133*	-0.0037	0.0144**	-0.0058	0.0162**
	(0.0040)	(0.0071)	(0.0038)	(0.0066)	(0.0040)	(0.0065)
internet	0.1710**	0.2999*	0.1449**	0.2362	0.1603**	0.2180
	(0.0689)	(0.1674)	(0.0735)	(0.1673)	(0.0726)	(0.1617)
_cons	2.7041***	1.0883	2.5307**	0.8208	2.5319**	0.8128
	(1.0264)	(1.7793)	(1.0971)	(1.8637)	(1.0977)	(1.8656)
N	4941	2880	4327	2773	4327	2773
pseudo R^2	0.122	0.120	0.121	0.122	0.122	0.123

Panel B: Replace distance with time spent

Sample	Urban areas	Rural areas	Urban areas	Rural areas	Urban areas	Rural areas
DIS=	*ti_live*	*ti_live*	*ti_aver*	*ti_aver*	*ti_min*	*ti_min*
DIS	-0.0034**	-0.0040***	-0.0026**	-0.0040***	-0.0026**	-0.0040***
	(0.0016)	(0.0010)	(0.0012)	(0.0010)	(0.0013)	(0.0010)
*internet*DIS*	-0.0014	0.0054*	-0.0006	0.0079***	-0.0010	0.0080***
	(0.0024)	(0.0028)	(0.0020)	(0.0028)	(0.0021)	(0.0027)
internet	0.1604**	0.2708*	0.1184	0.1128	0.1244	0.1345
	(0.0761)	(0.1603)	(0.0791)	(0.1670)	(0.0776)	(0.1602)
_cons	2.7548***	1.6145	2.7238**	1.3860	2.7765***	1.2664
	(1.0064)	(1.7285)	(1.0798)	(1.8115)	(1.0777)	(1.8134)
N	4996	3097	4362	2968	4389	2975
pseudo R^2	0.121	0.124	0.119	0.122	0.120	0.126

Note. the results of control variables are omitted for brevity.

6.2 The Role of Non-Internet Information Media

Next, we will investigate other information media to examine whether the role of the Internet can be replaced by other information media. According to the respondents' choice of other access to information besides Internet, we construct two dummy indicators: *telev* and *radio*. The former equals to 1 if the respondents use TV to obtain information and 0 otherwise. The latter equals 1 if the respondents use radio to obtain information and 0 otherwise. We use these two variables to replace the Internet in 6, to see whether they can play the same role as the Internet. The results are shown in Table 13.

Panel A and Panel B are the results of using television instead of the Internet and using radio instead of the Internet to obtain information, respectively. The results show that for all three distance indicators, non-internet information media can't alleviate the negative effect of distance on family entrepreneurship, indicating that the role of Internet can't be replaced by other information media.

Table 13. The Role of non-Internet Information Media

	(1)	(2)	(3)	(4)	(5)	(6)
Variables	*selfemp*	*selfemp*	*selfemp*	*selfemp*	*selfemp*	*selfemp*
DIS=	*dis_live*	*dis_live*	*dis_aver*	*dis_aver*	*dis_min*	*dis_min*
Sample	Urban areas	Rural areas	Urban areas	Rural areas	Urban areas	Rural areas
Panel A TV						
DIS	-0.0056	-0.0034	-0.0041	0.0003	-0.0064	0.0006
	(0.0040)	(0.0045)	(0.0038)	(0.0040)	(0.0041)	(0.0040)
*telev*DIS*	0.0012	-0.0021	0.0009	-0.0059	0.0030	-0.0060
	(0.0043)	(0.0047)	(0.0040)	(0.0043)	(0.0043)	(0.0043)
telev	-0.2129**	0.0521	-0.2571***	0.1190	-0.2776***	0.1215
	(0.0877)	(0.1629)	(0.0970)	(0.1656)	(0.0958)	(0.1642)
_cons	2.6087***	1.4310	2.5438**	1.1952	2.5952**	1.0977
	(1.0080)	(1.7194)	(1.0809)	(1.8037)	(1.0797)	(1.8030)
N	4987	3070	4360	2943	4375	2949
pseudo R^2	0.121	0.111	0.120	0.114	0.121	0.114
Panel B Radio						
DIS	-0.0039**	-0.0054***	-0.0027*	-0.0046***	-0.0030*	-0.0045***
	(0.0018)	(0.0017)	(0.0016)	(0.0016)	(0.0017)	(0.0016)
*radio*DIS*	-0.0101*	0.0030	-0.0084*	-0.0003	-0.0102*	-0.0000
	(0.0055)	(0.0062)	(0.0050)	(0.0066)	(0.0056)	(0.0064)
radio	0.1349	-0.2572	0.1481	-0.2025	0.1510	-0.2087
	(0.1060)	(0.2274)	(0.1113)	(0.2343)	(0.1098)	(0.2313)
_cons	2.5749**	1.4044	2.5030**	1.2050	2.5291**	1.1221
	(1.0129)	(1.7023)	(1.0843)	(1.7831)	(1.0840)	(1.7816)
N	4987	3070	4360	2943	4375	2949
pseudoR^2	0.119	0.112	0.118	0.114	0.119	0.114

Note. the results of control variables are omitted for brevity.

7. Conclusions

The high degree of dependence of traditional economic activities on spatial location has made "distance" an important factor affecting economic development (Henderson, 1974; Krugman, 1991; Anderson, 2011). However, the development of the Internet has realized the time-free transmission of information, reducing the adverse effects caused by geographic distance and offering more opportunities for economic activities. Existing research has investigated the role of the Internet at the international level (Freund & Weinhold, 2004; Gnangnon & Iyer, 2018). As a supplement to the micro level, this paper takes family entrepreneurship as an example to explore the impact of the Internet in moderating the adverse effects of distance, which not only enriches the existing research on the theme of the Internet, geographic distance and economic activities, but also demonstrates a marginal advance on the value creation mechanism of the Internet in promoting family entrepreneurship. In addition, this paper emphasizes the importance of the Internet in rural areas, providing some enlightenment for the development of transition economies and/or less-developed countries.

This paper shows that the Internet will have a positive impact on family entrepreneurship by at least partially compensating for the obstacles of geographic distance. We estimate the distance between the family residence and city center, and test the hypotheses empirically, using data from the China Household Finance Survey (CHFS). The results indicate that: first, the farther away the family residence is from the city center, the less likely the family is to carry out entrepreneurial activities; second, the Internet can moderate the negative impact of distance on family entrepreneurship, but this effect is not significant among urban families; third, the Internet helps to reduce the urban-rural differences of family entrepreneurship; furthermore, economic and social information obtained from the Internet can play a significant role while entertainment information does not. Our conclusions are robust after considering sample adjustment, an instrumental variable, and doing a PSM test.

The conclusions of this paper directly confirm the importance of the Internet, especially in rural areas. We believe that transition economies should not only continue to vigorously support the development of Internet technology, but also guide its application in remote areas, in order to give full play to the value creation function of the Internet for remote and underprivileged areas.

Acknowledgments: This research was supported by the Guangdong Provincial Philosophy and Social Sciences Discipline Co construction Project (GD24XYJ07).

References

Acs, Z. J., & Armington, C. (2004). Employment Growth and Entrepreneurial Activity in Cities. *Regional Studies*, *38*(8), 911-927.

Acs, Z. J., & Audretsch, D. B. (1990). *Innovation and Small Firms*. Cambridge Mass: MIT Press.

Acs, Z. J., & Varga, A. (2005). Entrepreneurship, Agglomeration and Technological Change. *Small Business Economics*, *24*(3), 323-334.

Andersson, A. (1985). "Creativity and Regional Development". *Papers of the Regional Science Association*, *56*(1), 5-20.

Anderson, J. E. (2011). The Gravity Model. *Annual Review of Economics*, *3*(1), 133-160.

Arundel, A., & Geuna, A. (2004). Proximity and the Use of Public Science by Innovative European Firms. *Economics of Innovation & New Technology*, *13*(6), 559-580.

Barry, L. (2011). Knowledge-Based Economic Activity and Regional Dynamics: Convergence or Divergence in Regional Income Levels and Implications for Labour Markets? *International Conference: New Socio-Economic Challenges of Development in Europe 2008, Labour Market Issues*, 21-30.

Birley, S. (1985). The Role of Networks in the Entrepreneurial Process. *Journal of Business Venturing*, *1*(1), 107-117.

Dabson, B. (2001). Supporting Rural Entrepreneurship. In *Rural and Agricultural Conferences Proceedings* (pp. 35-47).

Démurger, A., M. Gurgand, S. Li., & Yue, X. (2009). "Migrants as Second-Class Workers in Urban China? A Decomposition Analysis". *Journal of Comparative Economics*, *37*(4), 610-628.

Dettling, L. J. (2017). Broadband in the Labor Market: The Impact of Residential High-Speed Internet on Married Women's Labor Force Participation. *Industrial and Labor Relations Review*, *70*(2), 451-482.

Forero, B. (2013). Mobile Communication Networks and Internet Technologies as Drivers of Technical Efficiency Improvement. *Information Economics & Policy*, *25*(3), 126-141.

Forman, C., Goldfarb, A., & Greenstein, S. (2012). The Internet and Local Wages: A Puzzle. *American Economic Review*, *102*(l), 556-575.

Freund, C., & Weinhold, D. (2004). The Effect of the Internet on International Trade. *Journal of International Economics*, *62*(1), 171-189.

Fritsch, M., & Weyh, A. (2006). How Large Are the Direct Employment Effects of New Businesses? An Empirical Investigation for West Germany. *Small Business Economics*, *27*, 245-260.

Fujita, M., Krugman, P., & Venables, A. (1999). *The Spatial Economy: Cities, Regions and International Trade*. Cambridge Mass: MIT Press.

Fujita, M., & Rivera-Batiz, F. L. (1988). Agglomeration and Heterogeneity in Space: Introduction. *Regional Science & Urban Economics*, *18*(1), 1-5.

Gassmann, O., & Keupp, M. M. (2007). The Competitive Advantage of Early and Rapidly Internationalising SMEs in the Biotechnology Industry: A Knowledge-based View. *Journal of World Business*, *42*(3), 350-366.

Glaeser, E. L. (2007). Entrepreneurship and the City. NBER Working Paper, *13551*.

Glaeser, E. L., Kolko, J., & Saiz, A. (2000). "Consumer City". *NBER Working Paper*, 7790.

Granovetter, M. S. (1977). The Strength of Weak Ties. *Social Networks*, *78*(6), 347-367.

Gnangnon, S. K., & Iyer, H. (2018). Does Bridging the Internet Access Divide Contribute to Enhancing

Countries' Integration into the Global Trade in Services Markets? *Telecommunications Policy*, *42*(1), 61-77.

Guillen, M. F., & Suarez, S. L. (2001). "Developing the Internet: Entrepreneurship and Public Policy in Ireland, Singapore, Argentina, and Spain", *Telecommunications Policy*, *25*(5), 349-371.

Henderson, J. R., 2002, "Are High-growth Entrepreneurs Building the Rural Economy?", *Main Street Economist*, August, pp. 1-4.

Henderson, J. V. (1974). "The Sizes and Types of Cities", *American Economic Review*, *64*(4), 640-656.

Johansson, B., 1987, "Information Technology and The Viability of Spatial Networks". *Papers in Regional Science*, *61*(1), 51-64.

Krugman, P. (1991). Increasing Returns and Economic Geography. *The Journal of Political Economy*, *99*(3), *3*, 483-499.

Kuhn, P., & Mansour, H. (2014). Is Internet Job Search Still Ineffective?. *The Economic Journal*, *124*(581), 1213-1233.

Lamb, W. B., & Sherman, H. D. (2010). "Developing High-Growth Businesses in Rural Areas: A Study of Four U.S. States. *New England Journal of Entrepreneurship*, *13*(2), 9-19.

Lopez, M. V. (2009). Social Interactions and Information Dynamics in Self-employment in Mexico City. *Investigacion Economica*, *68*(268), 115-158.

Malecki, E. J., 1994, "Entrepreneurship in Regional and Local Development", *International Regional Science Review*, *16*, 119-153.

Paunov, C., & Rollo, V. (2016). "Has the Internet Fostered Inclusive Innovation in the Developing World?", *World Development*, *78*, 587-609.

Piore, M. J., & Sabel, C. F. (1986). *The Second Industrial Divide: Possibilities for Prosperity*, Published by Basic Books.

Psaltopoulos, D., S. Stathopoulou and D. Skuras, 2005, "The Location of Markets, Perceived Entrepreneurial Risk, and Start-up Capital of Micro Rural Firms", *Small Business Economics*, *25*(2), pp.147-158.

Rosenthal, S. S. and W. C. Strange, 2005, "The Geography of Entrepreneurship in the New York Metropolitan Area", *Economic Policy Review*, December, pp. 29-53.

Shen, K. N., V. Lindsay and Y. Xu, 2018, "Digital Entrepreneurship", *Information Systems Journal*, 28(6), pp. 1125-1128.

Song, Y., 2015, "From Offline Social Networks to Online Social Networks: Changes in Entrepreneurship", *Informatica Economica*, *19*(2), 120-133.

Stam, E., 2010, "Entrepreneurship, evolution and geography", in The Handbook of Evolutionary Economic Geography, Eds. by Boschma, R. and R. Martin R., 307-348.

Tabuchi, T. (1998). Urban Agglomeration and Dispersion: A Synthesis of Alonso and Krugman. *Journal of Urban Economics*, *44*(3), 333-351.

Vemuri, V. K., & Siddiqi, S. (2009). Impact of Commercialization of the Internet on International

Trade: A Panel Study Using the Extended Gravity Model. *International Trade Journal*, *23*(4), 458-484.

Wennekers, S., Stel, A., Thurik, A. R., & Reynolds, P. (2005). Nascent Entrepreneurship and the Level of Economic Development. *Small Business Economics*, *24*(3), 293-309.

Wynne, C., Berthon, P., Pitt, L., Ewing, M., & Napoli, J. (2001). The Impact of the Internet on the Distribution Value Chain: The Case of the South African, Tourism Industry. *International Marketing Review*, *18*(4), 420-431.

Zaheer, S., & Manrakhan, S. (2001). Concentration and Dispersion in Global Industries: Remote Electronic Access and the Location of Economic Activities. *Journal of International Business Studies*, *32*(4), 667-686.

Notes

Note 1. According to the definition of Ali Research Institute, Taobao village refers to a village where the number of active online stores reaches more than 10% of local households and the annual transaction volume of e-commerce reaches more than 10 million yuan. The first batch of Taobao villages discovered in 2009 include qingyanliu village, Yiwu City, Jinhua City, Zhejiang Province, Dongfeng Village, Suining County, Xuzhou City, Jiangsu Province, and donggaozhuang village, Qinghe County, Xingtai City, Hebei Province.

Note 2. CCTV News introduced this on December 15, 2018. Please refer to CNTV for details: http://tv.cntv.cn/video/C10437/09c8dcb499534c2aaf499ce2428bc876.

Note 3. The first Taobao village was identified in 2009, when the formation level of large-scale production organizations within families was relatively low. We chose to use the data of 2011 for analysis in order to better explore the mechanism of Internet penetration in remote areas.

Note 4. Including the average distance between the residence and the city center at 19.47km, the average distance between family owned houses and the city center at 20.98km, and the average minimum distance from the city center at 20.38km.

Note 5. We judged it according to the answer of the question in the questionnaire, "if you have an asset, what kind of investment project would you like to choose?" If "high risk, high return project" or "slightly high risk, slightly high return project" is selected, take 1; if "average risk, average return project" is selected, take 0; if "slightly low risk, slightly low return project" or "unwilling to bear any risk" is selected, take - 1.

Note 6. That is to say, some respondents have a rural residence registration, but live in urban areas.